CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Edinburgh Research Explorer

THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

The Bjorken sum rule with Monte Carlo and Neural Network techniques

Citation for published version:
Del Debbio, L, Guffanti, A & Piccione, A 2009, 'The Bjorken sum rule with Monte Carlo and Neural Network techniques', *Journal of High Energy Physics*, vol. 2009, no. 11, 060. https://doi.org/10.1088/1126-6708/2009/11/060

Digital Object Identifier (DOI):
10.1088/1126-6708/2009/11/060

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Early version, also known as pre-print

Published In:
Journal of High Energy Physics

Publisher Rights Statement:
Author's Post-print: author can archive post-print (ie final draft post-refereeing)

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jan. 2021

Preprint typeset in JHEP style - HYPER VERSION FREIBURG-PHENO-09/04

The Bjorken sum rule with Monte Carlo and Neural Network techniques

arXiv:0907.2506v1 [hep-ph] 15 Jul 2009

L. Del Debbio
SUPA, School of Physics and Astronomy, University of Edinburgh
Edinburgh EH9 3JZ, Scotland
Email: luigi.del.debbio@ed.ac.uk

A. Guffanti
Physikalisches Institut, Albert-Ludwigs-Univerität
Hermann-Herder-Straße 3, 79104 Freiburg, Germany
Email: alberto.guffanti@physik.uni-freiburg.de

A. Piccione
I.T.I.S. Pininfarina
Via Ponchielli 16, 10024 Moncalieri, Italy
and
I.N.F.N. Milano
Via Celoria 16, 20133 Milano, Italy
Email: hep@andreapiccione.it

Abstract: Determinations of structure functions and parton distribution functions have been recently obtained using Monte Carlo methods and neural networks as universal, unbiased interpolants for the unknown functional dependence. In this work the same methods are applied to obtain a parametrization of polarized Deep Inelastic Scattering (DIS) structure functions. The Monte Carlo approach provides a bias–free determination of the probability measure in the space of structure functions, while retaining all the information on experimental errors and correlations. In particular the error on the data is propagated into an error on the structure functions that has a clear statistical meaning. We present the application of this method to the parametrization from polarized DIS data of the photon asymmetries A_1^p and A_1^d from which we determine the structure functions $g_1^p(x,Q^2)$ and $g_1^d(x,Q^2)$, and discuss the possibility to extract physical parameters from these parametrizations. This work can be used as a starting point for the determination of polarized parton distributions.

Keywords: Polarized DIS, structure functions, g_1^p, Neural Networks.

Contents

1. Introduction

In QCD the description of scattering processes at large momentum transfer ($Q^2 \gg \Lambda^2_{QCD}$) involving (polarized) hadrons in the initial state is based on the factorization theorem. The latter allows a separation between the high–energy dynamics, described by coefficient functions which are calculable in perturbative QCD, from low–energy, non-perturbative effects, binding partons into hadrons, which are encoded into (polarized) parton distribution functions (PDFs).

The growth in statistics and increase in precision of data from experiments involving polarized hadrons scattering calls for a more accurate determination of polarized PDFs and their errors. A crucial problem in this respect is the determination of the uncertainty on a function (i.e. a probability measure on a space of functions) from a finite set of experimental data points. In the standard PDF extraction approach to the problem the infinite–dimensional space of continuous functions is mapped into a finite–dimensional space of parameters by choosing a particular basis in the space of functions and truncating the basis to a finite number of elements. This procedure entails some degree of arbitrariness. Any sensible choice must strike a balance between two competing requirements: on the one hand a small number of parameters introduces a bias in the determination of both

the functional form and the errors, as the chosen parametrization would not allow enough flexibility; on the other hand a large number of parameters could spoil the convergence of the fit, or be too sensitive to the statistical fluctuation of the experimental data.

This problem has been addressed by the NNPDF Collaboration in the case of unpolarized Deep Inelastic Scattering (DIS) structure functions in Refs. [1, 2], and in the case of the unpolarized PDFs in Refs. [3–5] using a method based on statistical inference and neural networks as an interpolating tool.

While avoiding technical complications linked to the extraction of PDFs from observables, the determination of structure functions addresses the main issue of devising a faithful estimation of errors on a function extracted from experimental data. The main ingredient in the studies above is the usage of Monte Carlo methods to obtain a representation of the probability measure in the space of structure functions. An ensemble of artificial data is generated, which reproduces all the statistical features (i.e. variances and correlations) of the original experimental data. Each set of artificial data is called a replica. A structure function, parametrized by a neural network, is then fitted to each replica. The net result of this procedure is an ensemble of fitted functions. This ensemble of fitted functions provides a representation of the measure in the space of structure functions. Errors and correlations of any observable involving the structure functions are obtained averaging over the ensemble of fits. Moreover suitable statistical estimators can be defined from the Monte Carlo ensemble which provide a quantitative description of the possible biases and inconsistencies in the fitting procedure. This method has been described in great detail in Refs. [1–4] to which the interested reader should refer.

The aim of this work is to apply the same techniques to obtain a bias–free parametrization of the photon asymmetries A_1^p and A_1^d from available polarized DIS data and extract from them the corresponding structure functions g_1^p and g_1^d. We provide further testing of the Monte Carlo method, and produce statistically meaningful error bars for the structure function. Besides allowing us to address all systematics related to the data and the method, such a parametrization might be an ideal input for a fit based on factorization scheme-invariant evolution equations to determine α_s, as proposed in Refs. [6, 7]. As shown in this work, a careful treatment of statistical and systematic errors leads to a reliable extraction of physically meaningful parameters such as α_s, g_A, and the higher–twist contributions to the structure functions. While these are not the best determinations available for these parameters, the results we obtain are in agreement with other determinations, and show the robustness of the Monte Carlo method.

We shall now discuss in turn the two steps that are needed to produce the Monte Carlo sample of fitted functions: first the treatment of the experimental data, and then the actual fitting procedure. The experimental data points included in the fit are discussed in Section 2; Section 3 summarizes briefly the NNPDF approach and the characteristics of the neural networks used for this particular study. The results of our fits, together with their phenomenological implications are presented and discussed in Sections 4 and 5.

2. Experimental data

The cross section asymmetry for parallel and anti-parallel configurations of longitudinal beam and target polarizations is given by:

$$A_{||} = \frac{\sigma^{\uparrow\downarrow} - \sigma^{\uparrow\uparrow}}{\sigma^{\uparrow\downarrow} + \sigma^{\uparrow\uparrow}} \tag{2.1}$$

and it is related to the virtual–photon asymmetries A_1, A_2 by:

$$A_{||} = D(A_1 + \eta A_2) \simeq D A_1 \,. \tag{2.2}$$

The photon depolarization factor D depends on kinematic factors and on the ratio:

$$R(x, Q^2) = \frac{\sigma_L(x, Q^2)}{\sigma_T(x, Q^2)} \,, \tag{2.3}$$

where σ_L and σ_T are the longitudinal and the transverse cross sections respectively (see e.g. Refs. [8, 9] for a detailed definition of all the quantities).

The polarized structure functions g_1 and g_2 are related to the virtual-photon asymmetries by:

$$A_1(x, Q^2) = \frac{g_1(x, Q^2) - \gamma^2 g_2(x, Q^2)}{F_1(x, Q^2)} \,, \tag{2.4}$$

$$A_2(x, Q^2) = \gamma \frac{g_1(x, Q^2) + g_2(x, Q^2)}{F_1(x, Q^2)} \,; \tag{2.5}$$

where

$$F_1(x, Q^2) = \frac{(1 + \gamma^2)}{2x[1 + R(x, Q^2)]} F_2(x, Q^2) \,, \tag{2.6}$$

is the unpolarized structure function, $\gamma^2 = \frac{4m^2x^2}{Q^2}$, and m denotes the nucleon mass.

The main features of each experimental data set used in the present analysis are summarized in Tab. 1, and their kinematical coverage of the (x, Q^2)-plane is shown in Fig. 1. We observe that the kinematical coverage of the available data is rather small, especially when compared to the one of the available unpolarized DIS data, thus we will have a sizable region of the kinematical plane in which the fit extrapolates the behaviour extracted from the region covered by data.

From Tab. 1 we infer that the systematic errors are on average one order of magnitude smaller than the statistical ones. This justifies the procedure of neglecting correlations for systematic errors and the procedure of summing errors in quadrature when computing the figure of merit (χ^2) to be minimized in the fitting procedure.

Finally we notice that E155 data have been corrected to yield A_1 by adding in Eq. (2.4) the g_2 contribution evaluated with the Wandzura–Wilczek relation and using the parametrization of g_1/F_1 given in Ref. [14]: this shift is also added as a source of uncertainty in the total error of the data set.

Experiment	x range	$Q^2(\mathrm{GeV}^2)$ range	N_{dat}	$\langle\sigma_{\mathrm{stat}}\rangle$	$\langle\sigma_{\mathrm{syst}}\rangle$	$\langle\sigma_{\mathrm{norm}}\rangle$	Type	Ref.
Proton								
EMC	0.015 - 0.466	3.5 - 29.5	10	0.077	0.024	0.028	A_1	[10]
SMC	0.001 - 0.480	0.3 - 58.0	15	0.026	0.003	0.012	A_1	[11]
SMC low-x	0.0001 - 0.121	0.02 - 23.1	15	0.033	0.002	0.006	A_1	[12]
E143	0.031 - 0.75	1.27 - 9.52	28	0.045	0.016	0.012	A_1	[13]
E155	0.015 - 0.75	1.22 - 34.72	24	0.043	0.018	0.026	g_1/F_1	[14]
HERMES06	0.0058 - 0.7311	0.26 - 14.29	45	0.126	0.019	0.017	A_1	[15]
Deuteron								
COMPASS	0.0051 - 0.474	1.18 - 47.5	12	0.034	0.017	0.011	A_1	[16]
SMC	0.001 - 0.480	0.3 - 58.0	15	0.032	0.003	0.006	A_1	[11]
SMC low-x	0.0001 - 0.121	0.02 - 23.1	15	0.069	0.005	0.005	A_1	[12]
E143	0.031 - 0.75	1.27 - 9.52	28	0.066	0.011	0.008	A_1	[13]
E155	0.015 - 0.75	1.22 - 34.72	24	0.091	0.009	0.011	g_1/F_1	[14]
HERMES06	0.0058 - 0.7311	0.26 - 14.29	45	0.089	0.007	0.009	A_1	[15]

Table 1: The proton and deuteron experimental data sets included in the present analysis. We show the kinematic range, the number of points, the average statistical, systematic and normalization uncertainty, and the measured observable.

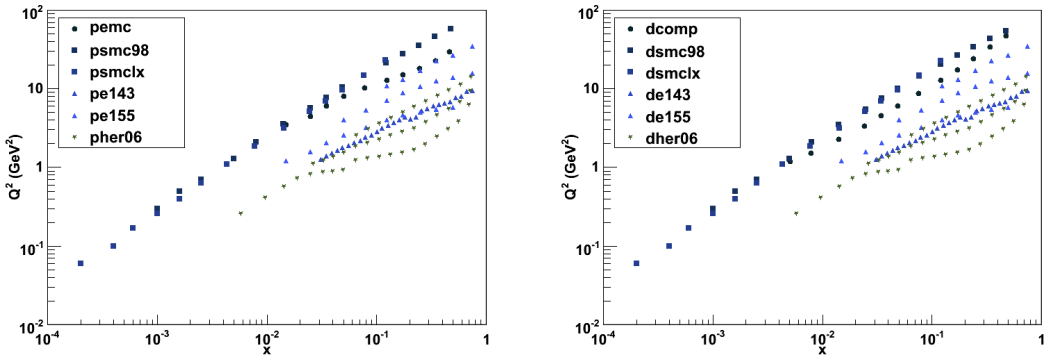

Figure 1: Experimental data in the (x, Q^2) plane used in the present analysis for the proton (left) and for the deuteron (right) target.

3. The NNPDF approach

In this section we briefly review the approach used to extract an unbiased determination of the asymmetry A_1 and the structure function g_1 from the available inclusive polarized DIS data, following the analysis performed by the NNPDF Collaboration for the determination of the unpolarized structure function F_2 [1, 2] and the parton densities [3–5].

The core idea underlying the NNPDF approach is based on using Monte Carlo methods to build a representation of the probability measure in the space of structure functions, and parametrizing the space of structure functions using neural networks. We refer the interested reader to the papers cited above for a detailed description of the methods and in the following we will briefly discuss the settings used in this analysis. It is worthwhile emphasizing that the Monte Carlo method does not require the use of neural networks,

and would yield a robust determination of the errors with any parametrization of the structure function, provided the parametrization is sufficiently flexible. A comparison of Monte Carlo analyses based on different parametrizations was performed in the framework of the HERA-LHC workshop by comparing the standard H1 and NNPDF analyses. The greater flexibility of the neural network parametrization compared to fixed functional forms is reflected in the larger error bands obtained using the NNPDF method, especially when considering the x region not covered by data (i.e. the extrapolation region). These features are illustrated by the results in Sects. 3.2 and 3.4 of Ref. [17].

3.1 Monte-Carlo replicas

We generate $N_{\rm rep}$ Monte-Carlo replicas of the experimental data according to

$$A_1^{({\rm art}),k}(x,Q^2) = \left(1 + r_{k,N}\frac{\sigma_N}{A_1^{({\rm exp})}(x,Q^2)}\right)\left[A_1^{({\rm exp})}(x,Q^2) + r_{k,t}\sigma_t(x,Q^2)\right], \tag{3.1}$$

where r_k are Gaussian distributed random numbers, σ_N is the quadratic sum of the normalization errors and σ_t is the total error, obtained by summing in quadrature the statistical and systematic errors, the latter assumed to be uncorrelated.

Following Ref. [18], the covariance matrix for experimental data points is evaluated using:

$$\mathrm{cov}_{ij} = \sigma_{N_i}\sigma_{N_j} + \delta_{ij}\sigma_{i,t}\,, \tag{3.2}$$

while during the fit for each replica, we minimize:

$$\chi^{2(k)} = \sum_{i=1}^{N_{\rm data}}\left(\frac{A_1^{(art),k}(x,Q^2) - A_1^{(net),k}(x,Q^2)}{\bar{\sigma}_{i,t}^{(k)}}\right), \tag{3.3}$$

where

$$\bar{\sigma}_{i,t}^{(k)} = \left(1 + r_{k,N}\frac{\sigma_N}{A_1^{(exp)}(x,Q^2)}\right)\sigma_{i,t}\,. \tag{3.4}$$

The number of Monte Carlo replicas of the data is determined by requiring that the average over the replicas reproduces the features (central values, errors and correlations) of the original experimental data to a required accuracy. The quantitative check is performed by means of the statistical estimators described in the appendix of Ref. [2] and the results for sets of 10, 100 and 1000 replicas are collected in Tab. 2 for the proton target data and in Tab. 3 for the deuteron target data. We observe that all the considered estimators have the correct scaling behaviour as the number of replica grows. We also point out that the large percentage error on the deuteron central values is due to a bulk of data whose values are close to zero.

3.2 Neural Networks as unbiased interpolants

Artificial neural networks, see e.g. Ref. [19], are a class of algorithms which provide a robust and universal approximant to incomplete or noisy data, with the only requirement of continuity. Neural networks are universal approximators for measurable functions [20].

	10	100	1000
$\langle PE\left[\langle A_1^{(art)}\rangle_{\rm rep}\right]\rangle$	14.20%	3.21%	1.83%
$r\left[A_1^{(art)}\right]$	0.974203	0.998114	0.999682
$\langle V\left[\sigma^{(art)}\right]\rangle_{\rm dat}$	$3.1\cdot10^{-3}$	$1.5\cdot10^{-3}$	$6.5\cdot10^{-4}$
$\langle PE\left[\langle\sigma^{(art)}\rangle\right]\rangle_{\rm dat}$	35.45%	12.44%	4.20%
$\langle\sigma^{(art)}\rangle_{\rm dat}$	0.0699	0.0766	0.0768
$r\left[\sigma^{(art)}\right]$	0.989956	0.997808	0.999793
$\langle V\left[\rho^{(art)}\right]\rangle_{\rm dat}$	$9.5\cdot10^{-2}$	$8.9\cdot10^{-3}$	$9.2\cdot10^{-4}$
$\langle\rho^{(art)}\rangle_{\rm dat}$	0.1469	0.1567	0.1585
$r\left[\rho^{(art)}\right]$	0.638102	0.944682	0.993523
$\langle V\left[{\rm cov}^{(art)}\right]\rangle_{\rm dat}$	$7.6\cdot10^{-5}$	$1.1\cdot10^{-5}$	$1.5\cdot10^{-6}$
$\langle{\rm cov}^{(art)}\rangle_{\rm dat}$	0.00166	0.00154	0.00160
$r\left[{\rm cov}^{(art)}\right]$	0.898803	0.986219	0.998858

Table 2: Statistical estimators for Monte Carlo replicas of A_1 for the proton data. The experimental data have $\langle\sigma^{(\rm exp)}\rangle_{\rm dat}=0.0764$, $\langle\rho^{(\rm exp)}\rangle_{\rm dat}=0.1566$, and $\langle{\rm cov}^{(\rm exp)}\rangle_{\rm dat}=0.00153$.

	10	100	1000
$\langle PE\left[\langle A_1^{(art)}\rangle_{\rm rep}\right]\rangle$	89.71%	24.90%	5.97%
$r\left[A_1^{(art)}\right]$	0.977524	0.98633	0.999865
$\langle V\left[\sigma^{(art)}\right]\rangle_{\rm dat}$	$5.3\cdot10^{-3}$	$1.7\cdot10^{-3}$	$6.7\cdot10^{-4}$
$\langle PE\left[\langle\sigma^{(art)}\rangle\right]\rangle_{\rm dat}$	35.03%	11.75%	4.4%
$\langle\sigma^{(art)}\rangle_{\rm dat}$	0.0689	0.0739	0.0734
$r\left[\sigma^{(art)}\right]$	0.977501	0.997965	0.999705
$\langle V\left[\rho^{(art)}\right]\rangle_{\rm dat}$	$1.0\cdot10^{-2}$	$9.2\cdot10^{-3}$	$8.7\cdot10^{-4}$
$\langle\rho^{(art)}\rangle_{\rm dat}$	0.0878	0.0904	0.0861
$r\left[\rho^{(art)}\right]$	0.612155	0.932158	0.992952
$\langle V\left[{\rm cov}^{(art)}\right]\rangle_{\rm dat}$	$6.3\cdot10^{-5}$	$1.0\cdot10^{-5}$	$1.0\cdot10^{-6}$
$\langle{\rm cov}^{(art)}\rangle_{\rm dat}$	0.00133	0.00145	0.00134
$r\left[{\rm cov}^{(art)}\right]$	0.959275	0.995339	0.999479

Table 3: Statistical estimators for Monte Carlo replicas of A_1 for the deuteron data. The experimental data have $\langle\sigma^{(\rm exp)}\rangle_{\rm dat}=0.0733$, $\langle\rho^{(\rm exp)}\rangle_{\rm dat}=0.0862$, and $\langle{\rm cov}^{(\rm exp)}\rangle_{\rm dat}=0.00135$.

This means that any continuous function can be approximated to any degree of accuracy by a sufficiently large neural network with one hidden layer and non-linear neuron activation function.

One of the main reasons to use neural networks in place of any other redundant parametrization is the existence of efficient techniques for *training* them, i.e. determining the parameters of the network (thresholds and weights) so that it reproduces a given set of input-output data. Equivalently one could say that a sufficiently large neural network provides a description of the data which is largely free of functional bias.

The analysis presented here uses a class of neural networks known as multilayer feed-

forward perceptrons, trained using a genetic algorithm [21,22]. The networks we employed have one hidden layer and a 2-4-1 architecture, which gives us a total of 17 free parameters for each network to be determined during the training. The guidance principle in the choice of the network architecture to be used is that it should provide a redundant parametrization for the data to be fitted, i.e. the network should have enough flexibility to fit not only the underlying physical law but also the statistical fluctuations of the experimental data. This property is crucial in ensuring that the fit results are not biased by the specific parametrization. The lack of functional bias is established a posteriori by verifying that fits performed with networks with different architectures lead to statistically equivalent results. This is achieved using the statistical estimators introduced in the NNPDF Collaboration's studies; the results of these comparisons are presented and discussed later.

The training of the individual networks to the Monte Carlo replicas is performed by minimizing the figure of merit given in Eq. (3.3). Given the extensive size and complex structure of the parameter space (a neural network with n parameters, weights and thresholds has $2n!$ equivalent global minimum configurations), the most efficient training algorithm turns out to be a genetic algorithm. The details of the implementation are discussed in Ref. [3].

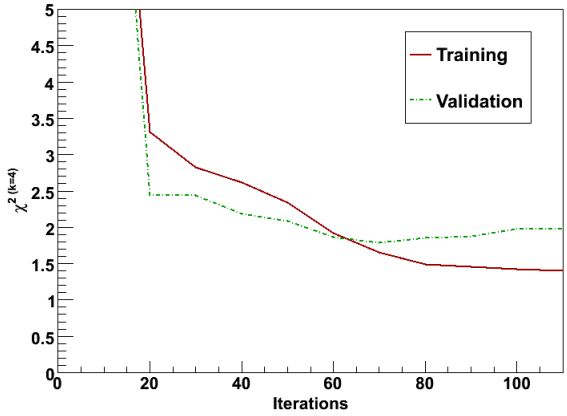

Figure 2: χ^2 for the training (red) and validation (green) sets of one replica in the reference fit to A_1^p.

As already pointed out in various references the fact that we adopt a redundant parametrization and that the figure of merit minimized in the training procedure is monotonically decreasing might lead to *overfitting* the data: the neural network reproduces not only the underlying physical law but also the statistical noise of the data sample. To prevent this from happening and to determine the optimal fit we adopt a criterion to stop our fit based on the *cross-validation* method. Once again our procedure is completely analogous to the one used for the unpolarized NNPDF fits.

For each replica of the experimental data we subdivide the data into a *training* and a *validation* set, respectively containing a fraction $f_{\rm tr}$ and $(1 - f_{\rm tr})$ of randomly chosen data points of each experiment.

We train one neural network on each replica of the data using the χ^2 of the training set as a figure of merit to be minimized. In parallel we compute the χ^2 of the validation set. We stop the training when we find that the χ^2 smeared over a given number of generations is decreasing for the training set while increasing for the validation set. A graphical illustration of such a behaviour for one of the replicas in the reference fit is given in Fig. 2.

4. Phenomenology

The study of the first moments of polarized structure functions is of phenomenological interest, since they can be used to extract information on the fraction of polarization carried by partons and on physical couplings. In the $\overline{\rm MS}$ scheme we have

$$\begin{aligned}\Gamma_1^{p,n}(Q^2) &= \int_0^1 dx\, g_1^{p,n}(x,Q^2) = \\ &= \frac{1}{36}\left[(a_8 \pm 3a_3)\Delta C_{\rm NS}^{\overline{\rm MS}}(\alpha_s(Q^2)) + 4a_0 \Delta C_{\rm S}^{\overline{\rm MS}}(\alpha_s(Q^2))\right] ,\end{aligned} \tag{4.1}$$

where $\Delta C_{\rm NS}^{\overline{\rm MS}}(\alpha_s(Q^2))$ and $\Delta C_{\rm S}^{\overline{\rm MS}}(\alpha_s(Q^2))$ are the first moments of the non-singlet and singlet Wilson coefficient functions, respectively, and

$$a_3 = (\Delta u + \Delta\bar{u}) - (\Delta d + \Delta\bar{d})\,, \tag{4.2}$$

$$a_8 = (\Delta u + \Delta\bar{u}) + (\Delta d + \Delta\bar{d}) - 2(\Delta s + \Delta\bar{s})\,, \tag{4.3}$$

$$a_0 = (\Delta u + \Delta\bar{u}) + (\Delta d + \Delta\bar{d}) + (\Delta s + \Delta\bar{s}) \equiv \Delta\Sigma\,. \tag{4.4}$$

Using isotopic spin invariance, it can be shown that a_3 is the axial coupling $g_A = G_A/G_V$ that governs neutron β-decay. Accurate measurements yield (see e.g. Ref. [9]):

$$g_A = 1.2670 \pm 0.0035\,. \tag{4.5}$$

The difference of the g_1 moments for proton and neutron leads to the Bjorken sum rule

$$\Gamma_1^{\rm NS}(Q^2) = \Gamma_1^p(Q^2) - \Gamma_1^n(Q^2) = \frac{1}{6} g_A \Delta C_{\rm NS}^{\overline{\rm MS}}(\alpha_s(Q^2)) + \delta_T + \delta_\tau\,; \tag{4.6}$$

where δ_T is the target mass correction and δ_t is the correction due to higher twists. Target mass corrections have been studied in Refs. [23–26] and can be evaluated for any moment n at the first order in m^2/Q^2 using [24]:

$$\delta_T = g_1^{(n)}(Q^2) - g_{10}^{(n)}(Q^2) = \frac{m^2}{Q^2}\frac{n(n+1)}{(n+2)^2}\left[(n+4) g_{10}^{n+2}(Q^2) + 4\frac{n+2}{n+1} g_{20}^{n+2}(Q^2)\right] , \tag{4.7}$$

where $g_i^{(n)}(Q^2) = \int_0^1 dx\, x^{n-1} g_i(x,Q^2)$ and g_{i0} is the structure function taken at zero mass of the nucleon. The higher–twist contribution is simply

$$\delta_\tau = \frac{\mu_4}{Q^2}\,, \tag{4.8}$$

where μ_4 can be extracted from experimental data at low Q^2 such as the CLAS data [27]. Finally, the coefficient function of Eq. (4.6) has been calculated in Ref. [28] and up to order α_s^3 is given by:

$$\begin{aligned}\Delta C_{\rm NS}^{\overline{\rm MS}}(\alpha_s(Q^2)) = 1 &- \frac{\alpha_s(Q^2)}{\pi} - \left(\frac{55}{12} - \frac{n_f}{3}\right)\left(\frac{\alpha_s(Q^2)}{\pi}\right)^2 \\ &- \left(41.4399 - 7.6072\, n_f + \frac{115}{648}\, n_f^2\right)\left(\frac{\alpha_s(Q^2)}{\pi}\right)^3 .\end{aligned} \tag{4.9}$$

For the running coupling we use the expanded solution of the renormalization group equation, up to NNLO we have:

$$\alpha_s\left(Q^2\right) = \alpha_s\left(Q^2\right)_{\rm LO}\left[1+\alpha_s\left(Q^2\right)_{\rm LO}\left[\alpha_s\left(Q^2\right)_{\rm LO}-\alpha_s\left(M_Z^2\right)\right](b_2-b_1^2)\right.$$
$$\left. +\ \alpha_s\left(Q^2\right)_{\rm NLO} b_1 \ln\frac{\alpha_s\left(Q^2\right)_{\rm NLO}}{\alpha_s\left(M_Z^2\right)}\right], \tag{4.10}$$

with

$$\alpha_s\left(Q^2\right)_{\rm NLO} = \alpha_s\left(Q^2\right)_{\rm LO}\left[1-b_1\alpha_s\left(Q^2\right)_{\rm LO}\ln\left(1+\beta_0\alpha_s\left(M_Z^2\right)\ln\frac{Q^2}{M_Z^2}\right)\right]\ , \tag{4.11}$$

$$\alpha_s\left(Q^2\right)_{\rm LO} = \frac{\alpha_s\left(M_Z^2\right)}{1+\beta_0\alpha_s\left(M_Z^2\right)\ln\frac{Q^2}{M_Z^2}}, \tag{4.12}$$

and the beta function coefficients given by

$$Q^2\frac{da_s(Q^2)}{dQ^2} = -\sum_{k=0}^{2}\beta_k a_s(Q^2)^{k+2}, \qquad a_s(Q^2)=\frac{\alpha_s(Q^2)}{4\pi}\,, \tag{4.13}$$

where

$$\begin{aligned}\beta_0 &= 11-\frac{2}{3}n_f\ , \\ \beta_1 &= 102-\frac{38}{3}n_f\ , \\ \beta_2 &= \frac{2857}{2}-\frac{5033}{18}n_f+\frac{325}{54}n_f^2\,,\end{aligned} \tag{4.14}$$

and $b_i \equiv \beta_i/\beta_0$.

5. Results

In this section we present our parametrization of the proton and deuteron asymmetries and the structure functions extracted from them.

We assess the quality of the fit by comparing our extraction with the experimental data included in the analysis, and by studying the stability of our results under variations of the parametrization used for the networks.

Then, as an example of a possible application of our result to a phenomenological analysis, we study the extraction of the physical parameters (the strong coupling constant α_s and the axial coupling g_A) from the Bjorken sum rule. In order to give a faithful error on the extracted quantities we study the impact of the different assumptions which are needed to reconstruct the structure functions and then to evaluate the Bjorken sum rule. Results are compared to existing estimates.

5.1 The final fit and its statistical features

In Tab. 4 we show the $\chi^2/N_{\rm data}$ for each target and each experimental data set included in the present analysis. We first observe the overall good quality of our fit. For the proton the small values of χ^2 for EMC, SMC and HERMES can be explained by a possible overestimate of experimental errors. For the deuteron all the χ^2 are of order 1, except for the E155 data set which has a value of χ^2 significantly smaller than one. The somewhat larger value of χ^2 for the E143 deuteron data set can be understood by looking at Figs. 3 and 4 where we present a comparison of our fit to experimental data in different kinematical regions. We observe that in the case of E143 the deuteron data show small incompatibilities among themselves, and the large value of χ^2 is a reflection of this. It is interesting to remark that a careful analysis of the χ^2 value for each experiment allows the identification of potential incompatible data. This feature had already been pointed out in the unpolarized studies by the NNPDF Collaboration.

Proton	χ^2	Deuteron	χ^2
EMC	0.370	COMPASS	0.885
SMC	0.480	SMC	1.100
SMC low-x	1.150	SMC low-x	0.774
E143	0.904	E143	1.530
E155	0.717	E155	0.661
HERMES06	0.456	HERMES06	0.881
Total	0.666	Total	0.986

Table 4: The χ^2 of the fit for proton and deuteron data.

In Tabs. 5, 6, and 7 we study the self–stability of the fit and the stability against the variation of the parametrization with respect to a smaller and a larger architecture. To this extent we define four different regions: one where we expect our fit to be an interpolation of the available data (Data region) and three where its behaviour is extrapolated to regions of the (x, Q^2)-plane not covered by present data:

- Data: $0.01 < x < 0.75$ and $2\ \text{GeV}^2 < Q^2 < 20\ \text{GeV}^2$;
- Low-x: $0.0001 < x < 0.001$ and $2\ \text{GeV}^2 < Q^2 < 20\ \text{GeV}^2$;
- Low-Q^2: $0.2 < x < 0.8$ and $0.1\ \text{GeV}^2 < Q^2 < 2\ \text{GeV}^2$;
- High-Q^2: $0.2 < x < 0.8$ and $20\ \text{GeV}^2 < Q^2 < 60\ \text{GeV}^2$.

We observe that all the estimators for self–stabilities are of order unity (or smaller), meaning that different subsets within the whole ensemble of replicas have the same statistical features.

When we compare our final fit to a fit performed using networks with a smaller architecture, we notice that the the two fits are statistically equivalent. The same happens for the comparison with a fit done with networks with a larger architecture, with the only exception of the errors on the deuteron fit in the extrapolation (all distances are order 1.5), which show some minor instability.

5.2 Structure functions reconstruction

In order to reconstruct the structure function g_1 from data on the asymmetry A_1 as given in Eq. (2.4) some additional assumptions are needed. In the following we assess the impact

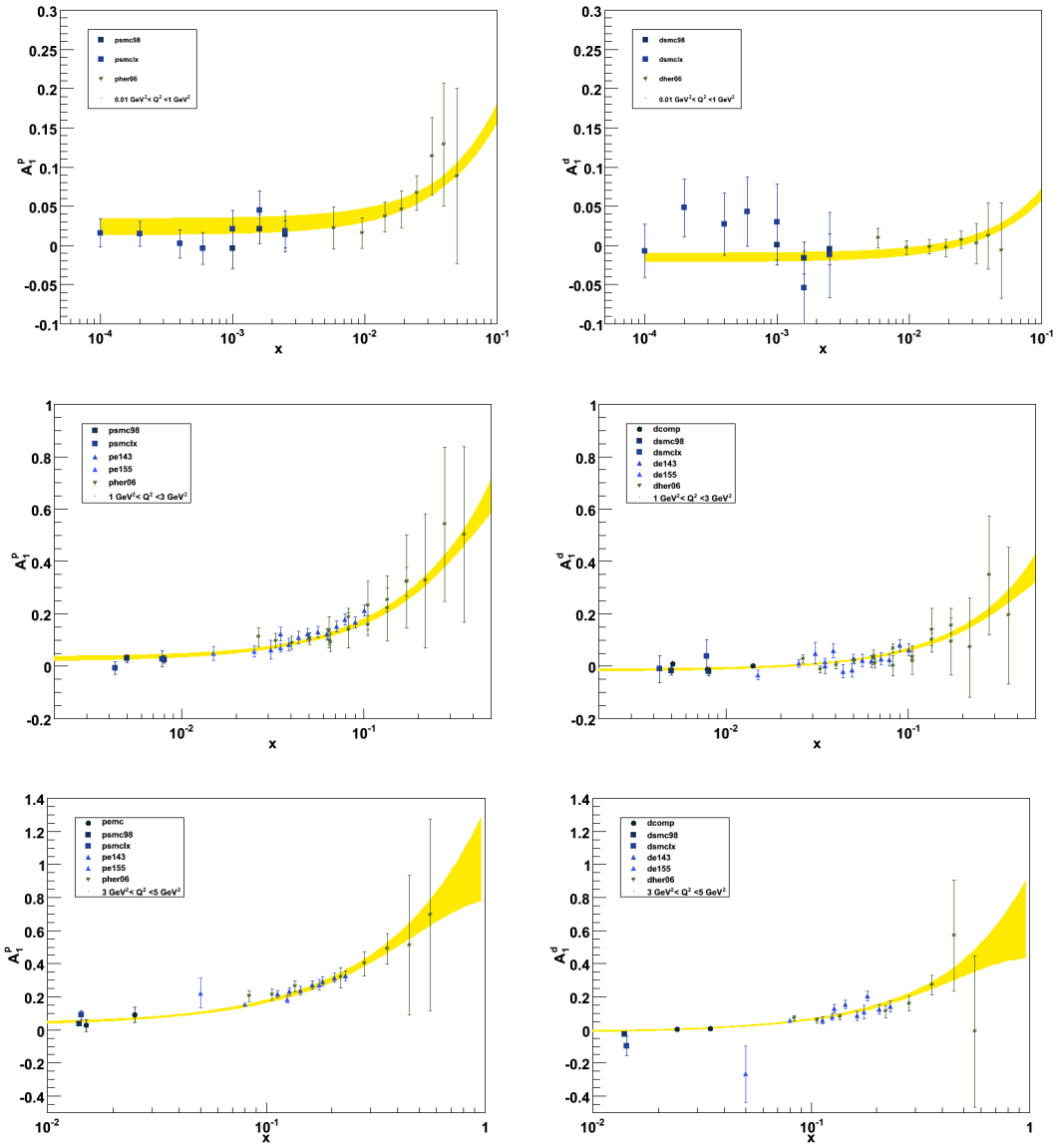

Figure 3: The fitted asymmetries compared to proton (left) and deuteron (right) data for $0.01\ \mathrm{GeV}^2 < Q^2 < 1\ \mathrm{GeV}^2$ (upper row), $1\ \mathrm{GeV}^2 < Q^2 < 3\ \mathrm{GeV}^2$ (central row) and $3\ \mathrm{GeV}^2 < Q^2 < 5\ \mathrm{GeV}^2$ (lower row). In the plots A_1 is evaluated at the central value of each Q^2 range.

of our assumptions for g_2, F_2 and R on the determination of first moment of g_1. These checks are done using an ensemble of 100 replicas, which is enough to this purpose, and in a range of x and Q^2 which is entirely in the data region in order to avoid any extrapolation effects. Finally, we compare our result for 1000 replicas with the sum rules obtained by experimental collaborations.

The first assumption whose impact we consider is the one on the structure function

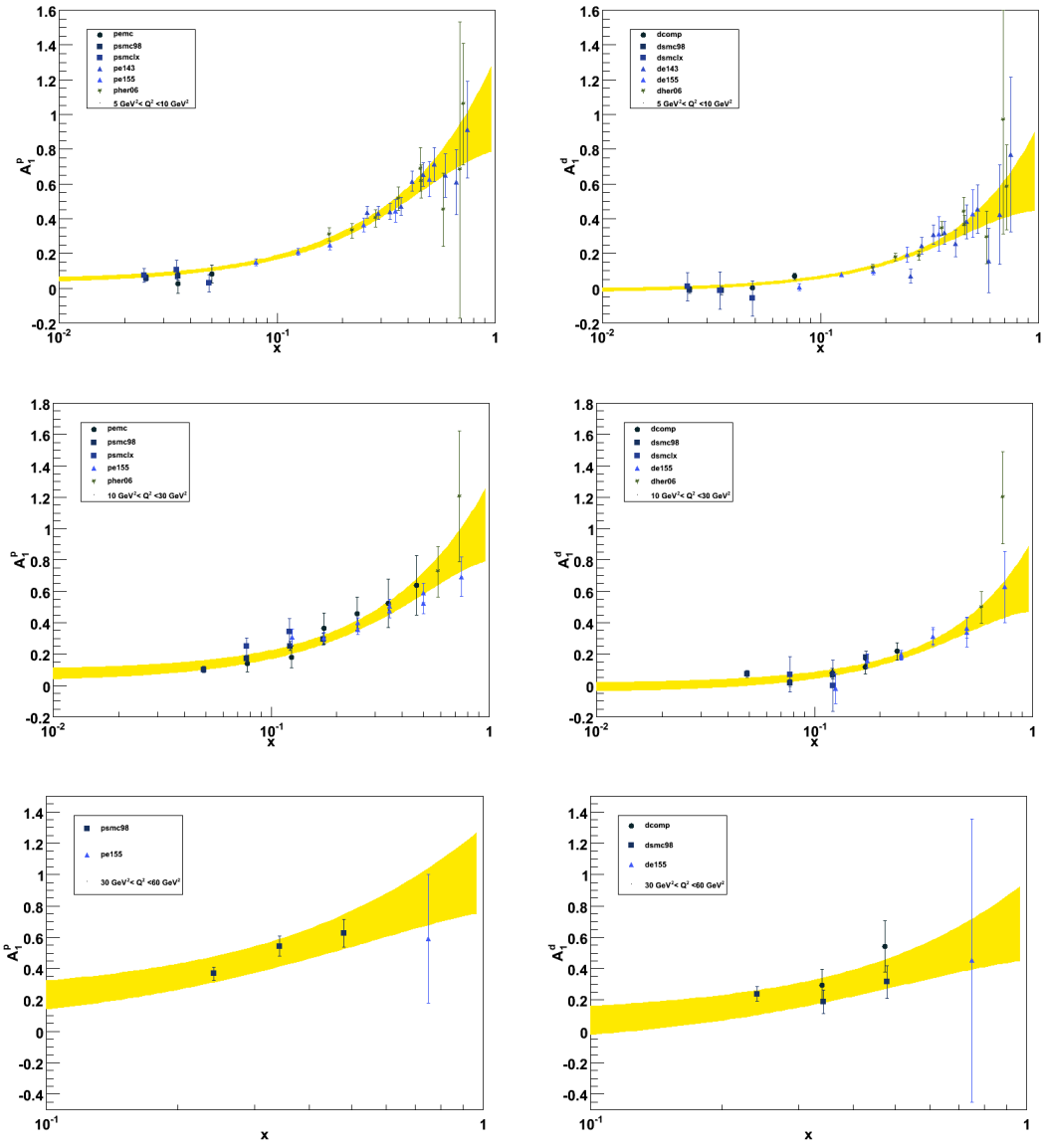

Figure 4: The fit compared to proton (left) and deuteron (right) data for 5 GeV$^2 < Q^2 <$ 10 GeV^2 (upper row), 10 GeV$^2 < Q^2 <$ 30 GeV^2 (central row) and 30 GeV$^2 < Q^2 <$ 60 GeV^2 (lower row). In the plots A_1 is evaluated at the central value of each Q^2 range.

g_2, which is evaluated from the Wandzura–Wilczek relation [29]

$$g_2^{WW}(x,Q^2) = -g_1(x,Q^2) + \int_x^1 \frac{dy}{y} g_1(y,Q^2)\,. \tag{5.1}$$

Inserting this expression into Eq. (2.4) gives

$$g_1(x,Q^2) = \frac{1}{1+\gamma^2}\left(A_1(x,Q^2)F_1(x,Q^2) + \gamma^2 \int_x^1 \frac{dy}{y} g_1(y,Q^2)\right)\,, \tag{5.2}$$

Proton	Data	Low-x	Low-Q^2	High-Q^2
$\langle d[A_1]\rangle$	0.963 ± 0.011	0.660 ± 0.011	1.085 ± 0.015	0.966 ± 0.014
$\langle d[\sigma_{A_1}]\rangle$	0.840 ± 0.006	0.618 ± 0.011	0.966 ± 0.014	0.905 ± 0.012
Deuteron	Data	Low-x	Low-Q^2	High-Q^2
$\langle d[A_1]\rangle$	0.772 ± 0.008	0.670 ± 0.012	0.804 ± 0.012	0.688 ± 0.011
$\langle d[\sigma_{A_1}]\rangle$	0.818 ± 0.007	0.899 ± 0.014	0.730 ± 0.011	0.773 ± 0.011

Table 5: Self–stability estimators evaluated with 100 replicas. The entries in the table show the statistical differences between results based on different subsets of 100 replicas randomly chosen in our Monte Carlo ensemble.

Proton	Data	Low-x	Low-Q^2	High-Q^2
$\langle d[A_1]\rangle$	0.952 ± 0.010	0.792 ± 0.014	0.859 ± 0.012	1.295 ± 0.016
$\langle d[\sigma_{A_1}]\rangle$	1.104 ± 0.008	1.405 ± 0.013	0.975 ± 0.014	1.002 ± 0.011
Deuteron	Data	Low-x	Low-Q^2	High-Q^2
$\langle d[A_1]\rangle$	1.217 ± 0.012	1.302 ± 0.012	1.324 ± 0.017	0.703 ± 0.010
$\langle d[\sigma_{A_1}]\rangle$	0.963 ± 0.008	1.199 ± 0.010	0.962 ± 0.010	1.689 ± 0.019

Table 6: Stability estimators for the reference fit (architecture 2-4-1) compared to a fit with a smaller architecture (2-3-1).

Proton	Data	Low-x	Low-Q^2	High-Q^2
$\langle d[A_1]\rangle$	1.076 ± 0.012	1.179 ± 0.017	0.693 ± 0.011	1.625 ± 0.018
$\langle d[\sigma_{A_1}]\rangle$	1.258 ± 0.010	1.354 ± 0.014	0.709 ± 0.010	1.151 ± 0.016
Deuteron	Data	Low-x	Low-Q^2	High-Q^2
$\langle d[A_1]\rangle$	0.867 ± 0.009	0.794 ± 0.012	0.856 ± 0.015	0.840 ± 0.017
$\langle d[\sigma_{A_1}]\rangle$	1.460 ± 0.009	1.941 ± 0.015	1.410 ± 0.011	1.303 ± 0.011

Table 7: Stability estimators for the reference fit (architecture 2-4-1) compared to a fit with a larger architecture (2-5-1).

which needs to be evaluated iteratively. To this purpose we take the initial value $g_1(x, Q^2)$ evaluated with $g_2(x, Q^2) = 0$, and we use

$$g_1^{(i_{WW})}(x, Q^2) = \frac{1}{1+\gamma^2}\left(A_1(x, Q^2)F_1(x, Q^2) + \gamma^2 \int_x^1 \frac{dy}{y} g_1^{(i_{WW}-1)}(y, Q^2)\right) . \quad (5.3)$$

From Tab. 8 we see that for Q^2 values above 2 GeV^2 one iteration is enough to stabilize the result for the first moment of g_1 computed in the data region. For lower scales, say $Q^2 \simeq 1$ GeV^2, at least two iterations of the Wandzura-Wilczek relation are needed in order to obtain a stable result. In the following the index i_{WW} will be omitted as the number of iterations used should be evident from the scale at which the first moment of g_1 is evaluated.

For the unpolarized structure function F_2 we use the parametrization given in Ref. [2] for the proton and the one given in Ref. [1] for the deuteron. Since these parametrizations have also been extracted using a Monte Carlo procedure, ensembles of replicas are available for F_2; hence the result for g_1 is evaluated as:

$$g_1(x, Q^2) = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} \left[A_1^{(k)}(x, Q^2) \frac{(1+\gamma^2)}{2x\,[1+R(x, Q^2)]} F_2^{(k)}(x, Q^2) + \gamma^2 g_2^{(k)}(x, Q^2)\right] , \quad (5.4)$$

Proton	$Q^2 = 2$ GeV^2	$Q^2 = 5$ GeV^2	$Q^2 = 10$ GeV^2	$Q^2 = 20$ GeV^2
$i_{ww} = 0$	0.1184 ± 0.0069	0.1196 ± 0.0068	0.1272 ± 0.0106	0.1401 ± 0.0216
$i_{ww} = 1$	0.1086 ± 0.0062	0.1154 ± 0.0066	0.1251 ± 0.0104	0.1391 ± 0.0215
$i_{ww} = 2$	0.1072 ± 0.0061	0.1152 ± 0.0065	0.1251 ± 0.0104	0.1391 ± 0.0215
Deuteron	$Q^2 = 2$ GeV^2	$Q^2 = 5$ GeV^2	$Q^2 = 10$ GeV^2	$Q^2 = 20$ GeV^2
$i_{WW} = 0$	0.0459 ± 0.0049	0.0414 ± 0.0036	0.0401 ± 0.0064	0.0397 ± 0.0149
$i_{WW} = 1$	0.0414 ± 0.0044	0.0396 ± 0.0035	0.0392 ± 0.0064	0.0393 ± 0.0149
$i_{WW} = 2$	0.0407 ± 0.0042	0.0395 ± 0.0034	0.0392 ± 0.0064	0.0393 ± 0.0149

Table 8: First moment for x between 0.01 and 0.75 for different number of iterations of the Wandzura–Wilczek relation.

which takes into account both the uncertainty on A_1 and the one on F_2 (with $g_2^{(k)}(x, Q^2)$ we denote the expression in Eq. (5.1) evaluated for the k-th replica). Since there is no correlation between the extraction of A_1 and the one of F_2 the replicas of A_1, and F_2 can be sampled independently.

In order to estimate the contribution of the uncertainty on F_2 to the uncertainty on g_1, we can recompute g_1 as

$$g_1(x, Q^2) = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} \left[A_1^{(k)}(x, Q^2) \frac{(1+\gamma^2)}{2x\,[1 + R(x, Q^2)]} \langle F_2 \rangle(x, Q^2) + \gamma^2 g_2^{(k)}(x, Q^2) \right] , \quad (5.5)$$

where for each k-th replica of A_1 we use the averaged value of the unpolarized structure function

$$\langle F_2 \rangle(x, Q^2) = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} F_2^{(k)}(x, Q^2) \, . \quad (5.6)$$

This procedure clearly freezes the fluctuations in F_2, which is kept fixed to its average value. The result is given in Tab. 9, where we see that the contribution to the uncertainty on the first moment of g_1 due to F_2 is negligible. In the following we will always use g_1 as given from Eq. (5.4).

Proton	$Q^2 = 2$ GeV^2	$Q^2 = 5$ GeV^2	$Q^2 = 20$ GeV^2
Eq. (5.4)	0.1086 ± 0.0062	0.1154 ± 0.0066	0.1391 ± 0.0215
Eq. (5.5)	0.1086 ± 0.0059	0.1154 ± 0.0064	0.1391 ± 0.0213
Deuteron	$Q^2 = 2$ GeV^2	$Q^2 = 5$ GeV^2	$Q^2 = 20$ GeV^2
Eq. (5.4)	0.0414 ± 0.0044	0.0396 ± 0.0035	0.0393 ± 0.0149
Eq. (5.5)	0.0414 ± 0.0038	0.0396 ± 0.0034	0.0394 ± 0.0150

Table 9: First moment for x between 0.01 and 0.75 with and without the error on F_2

Finally a parametrization of $R(x, Q^2)$ is needed in order to extract g_1 from A_1. Here we use $R_{SLAC}(x, Q^2)$ given in Ref. [30, 31]. Such a parametrization provides also an error estimate, which we use to assess the impact of $R_{SLAC}(x, Q^2)$ on the total uncertainty of the first moment of g_1. In Tab. 10 we compare the sum rule evaluated with the central value of $R_{SLAC}(x, Q^2)$ with the one obtained by taking into account the error on $R_{SLAC}(x, Q^2)$. This is achieved by letting $R_{SLAC}(x, Q^2)$ fluctuate within its own error in the Monte Carlo

sample; for the k-th replica we use:

$$R_{SLAC}(x,Q^2) + r^{(k)}\Delta R_{SLAC}(x,Q^2)\,, \tag{5.7}$$

where $\Delta R_{SLAC}(x,Q^2)$ is the error on the parametrization, and $r^{(k)}$ is a univariate Gaussian random number. Since $R_{SLAC}(x,Q^2)$ is a parametrization of experimental data, we take the error as a statistical one, with no correlation between different replicas, and thus we use a different random number each time a value of $R_{SLAC}(x,Q^2)$ is needed. From the results collected in Tab. 10 we conclude that the error on $R_{SLAC}(x,Q^2)$ is also negligible.

Proton	$Q^2 = 2$ GeV^2	$Q^2 = 5$ GeV^2	$Q^2 = 20$ GeV^2
$R_{SLAC}(x,Q^2)$	0.1086 ± 0.0062	0.1154 ± 0.0066	0.1391 ± 0.0215
$R_{SLAC}(x,Q^2) + r^{(k)}\Delta R_{SLAC}(x,Q^2)$	0.1086 ± 0.0062	0.1154 ± 0.0066	0.1392 ± 0.0216
Deuteron	$Q^2 = 2$ GeV^2	$Q^2 = 5$ GeV^2	$Q^2 = 20$ GeV^2
$R_{SLAC}(x,Q^2)$	0.0414 ± 0.0044	0.0396 ± 0.0035	0.0393 ± 0.0149
$R_{SLAC}(x,Q^2) + r^{(k)}\Delta R_{SLAC}(x,Q^2)$	0.0415 ± 0.0044	0.0397 ± 0.0035	0.0395 ± 0.0148

Table 10: First moment for x between 0.01 and 0.75 with and without the error on R

We will now compare our results for the integral of g_1 at different scales and over different x ranges obtained using our ensemble of 1000 replicas with those obtained by different experimental collaborations.

In Tab. 11 results for the proton and the deuteron sum rules are compared to the result of Ref. [11]. We observe that the results are compatible within errors, and that our evaluation has a larger error.

Target	SMC98	This Analysis
$Q^2 = 10\text{GeV}^2$		
p	0.131 ± 0.009	0.139 ± 0.015
d	0.037 ± 0.007	0.035 ± 0.011

Table 11: Comparison of the proton and deuteron sum rules $\left(\int_{0.003}^{0.7} dx g_1(x,Q^2)\right)$ as determined in the present analysis with the results obtained by the SMC collaboration [11].

In Tab. 12 we compare our result with the ones in Refs. [13], and [15]. First we notice that the errors are of the same size, while our central values are systematically smaller, with a significant difference for the proton at low Q^2. A substantial part of effect can be attributed to the different parametrization used for the unpolarized structure function. Indeed, if we evaluate the sum rule of E143 with the SMC98 F_2^p parametrization [11, 32, 33], at $Q^2 = 2$ GeV^2 we obtain 0.116 ± 0.008 which is less than one sigma away from the result in Refs. [13]; the same happens for the HERMES06 case, using the ALLM parametrization [34] at $Q^2 = 2.5$ GeV^2 we get 0.1188 ± 0.0073. This can be understood looking at Fig. 5 where we compare the different parametrizations for F_2^p used in the different analysis.

It is clear that, while the different F_2^p parametrizations agree in the kinematical region covered by experimental data, they differ significantly at low-Q^2 in the large-x region where there are no data and an extrapolation is needed. For the ALLM and the SMC98 parametrizations the large-x behaviour is determined by the chosen functional form; the NNPDF parametrization interpolates by continuity from the last experimental point to the kinematical constrain $F_2(x=1,Q^2) = 0$. The difference among the parametrizations

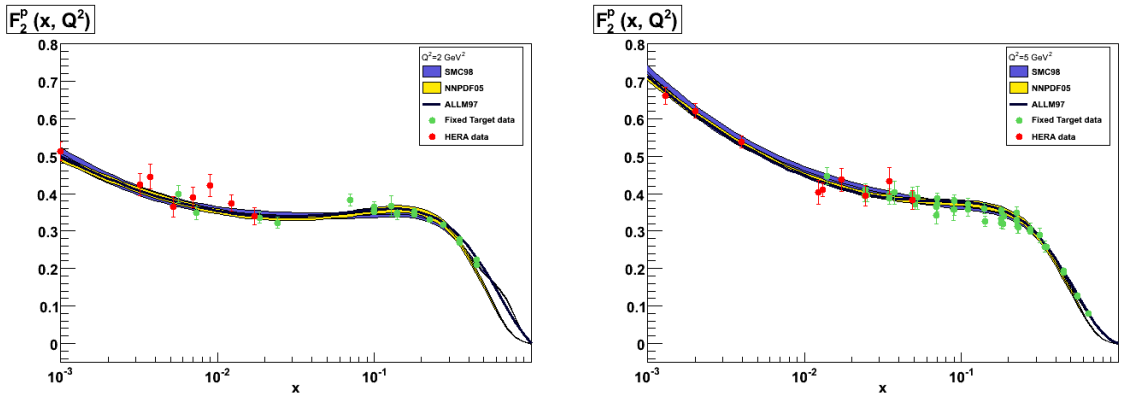

Figure 5: Comparison of NNPDF, SMC98 and ALLM parametrizations of the unpolarized structure function F_2^p in the region where we evaluate the Bjorken sum rule.

is then enhanced once we multiply by the asymmetry A_1 to reconstruct the polarized structure function g_1.

Target	E143	This Analysis
$Q^2 = 2\text{GeV}^2$		
p	0.120 ± 0.007	0.102 ± 0.007
d	0.047 ± 0.006	0.042 ± 0.005
n	-0.022 ± 0.013	-0.011 ± 0.011
NS	0.149 ± 0.016	0.113 ± 0.016
$Q^2 = 5\text{GeV}^2$		
p	0.116 ± 0.007	0.106 ± 0.006
d	0.043 ± 0.004	0.040 ± 0.003
n	-0.025 ± 0.009	-0.018 ± 0.009
NS	0.141 ± 0.013	0.124 ± 0.014

Target	HERMES	This Analysis
$Q^2 = 2.5\text{GeV}^2$		
p	0.1201 ± 0.0090	0.1055 ± 0.0066
d	0.0428 ± 0.0035	0.0416 ± 0.0043
n	-0.0276 ± 0.0093	-0.0154 ± 0.0107
NS	0.1477 ± 0.0167	0.1209 ± 0.0152
$Q^2 = 5\text{GeV}^2$		
p	0.1211 ± 0.0092	0.1097 ± 0.0065
d	0.0436 ± 0.0035	0.0407 ± 0.0033
n	-0.0268 ± 0.0094	-0.0218 ± 0.0093
NS	0.1479 ± 0.0169	0.1315 ± 0.0144

Table 12: Comparison of the integral of g_1 over different x ranges, at different scales, as determined form the present analysis with the results obtained form by E143, left pad: $\int_{0.03}^{0.8} dx g_1(x, Q^2)$ and HERMES, right pad: $\int_{0.021}^{0.9} dx g_1(x, Q^2)$

Since no neutron target data have been used in our fit, the neutron structure function, g_1^n, is evaluated from the proton and deuteron ones as

$$g_1^n(x, Q^2) = 2\frac{g_1^d(x, Q^2)}{1 - 1.5\omega_D} - g_1^p(x, Q^2)\,, \tag{5.8}$$

where ω_D is the probability of the deuteron to be in the D state; we use $\omega_D = 0.05$ which covers most of the published values [35]. We observe that, even if the neutron sum rule is a pure prediction, it is compatible with other estimations.

In Fig. 6 we show a comparison of the polarized structure function as extracted in this analysis to data, in the region where we evaluate the Bjorken sum rule. The comparison shows a good agreement.

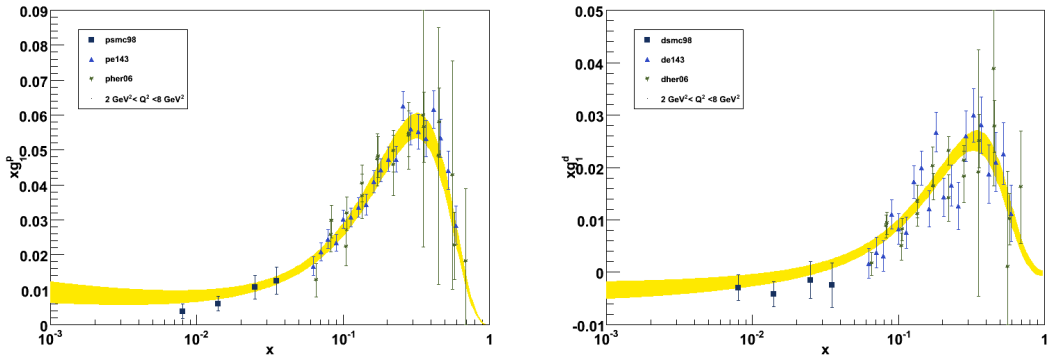

Figure 6: Plot of the structure functions in the region where we evaluate the Bjorken sum rule. The fit curves are taken at $Q^2 = 5$ GeV^2.

5.3 Extraction of couplings

In order to extract the strong coupling α_s and the axial coupling g_A from the values of the Bjorken sum rule we need to extrapolate our fit in the Bjorken variable x down to $x = 0$ and up to $x = 1$.

In this section we discuss the impact of these extrapolations on the extraction of the couplings, and we assess the impact of target mass corrections. Finally we present the results we obtain for α_s and g_A from our fit. All checks are performed with 100 replicas, while for final results we use the full set of 1000 replicas.

The extrapolation at large–x is embedded in the parametrization of F_2, as discussed in the previous section. Therefore we do not need any further assumption to constrain the large–x behaviour.

The low–x behavior of the structure function g_1 is instead very weakly constrained by data, and the Regge behaviour is usually assumed; following Ref. [36] we write:

$$g_1(x, Q^2) \simeq A\, x^b\,, \tag{5.9}$$

with $0 < b < 0.5$. Such an assumption requires to choose a value of $x_{\rm match}$ such that for $x < x_{\rm match}$ the Regge behaviour is assumed to set in. The normalization factor A in Eq. 5.9 is then determined by the matching condition

$$g_1^{\rm (fit)}(x_{\rm match}, Q^2) = A\, x_{\rm match}^b\,. \tag{5.10}$$

In order to choose the matching point, we fix the Regge exponent to 0.2 and then we proceed as in Ref. [37]: we evaluate the integrals at different values of Q^2 in the range $0 < x < 1$ for different choices of $x_{\rm match}$, and we look for the minimum value of the error for each value of Q^2.

From the results collected in Tab. 13 we see that $x_{\rm match}$ grows as Q^2 gets larger. This is understood looking at Fig. 1 where we see that for larger scales the coverage of the data moves towards higher values of x.

Once the matching point is been determined, in order to take into account the uncertainty on the value of the Regge exponent and the one on the choice of the matching point, we randomize the Regge exponent in the range $-0.1 < b < 0.5$ and we choose the matching point to be in the range $x_{\rm match} < x < 2x_{\rm match}$.

In Tab. 14 we present the comparison for the first moment of the structure function g_1 evaluated with and without the target mass correction as given in Eq. (4.7). We observe that the shift on the values of the moment due to the inclusion of these effects is smaller than the experimental error even at the lowest Q^2.

In principle we could extract g_A, α_s and the higher-twist term by fitting Eq. (4.6) evaluated from data at a given value of Q^2. In practice we evaluate N_{Q^2} different moments taken at different Q^2 in the kinematical region where we have a good coverage by experimental data: 2 GeV^2 $\leq Q^2 \leq$ 20 GeV^2. Indeed for $Q^2 >$ 20 GeV^2 the errors on the computed moments become so large that their weighted contribution in the combination is negligible. On the lower side of the energy range we choose to start from $Q^2 = 2$ GeV^2, since below this scale a perturbative QCD approach might not be reliable. For this reason we do not fit the higher-twist term, but we will access its contribution by varying the lower cut in Q^2.

Q^2(GeV^2)	$x_{\rm match}$	$\Gamma_1^{\rm NS}(Q^2)$	Error
1	0.0100	0.12499	0.020989
2	0.0100	0.1356	0.018239
3	0.0100	0.14324	0.017827
4	0.0200	0.13847	0.018275
5	0.0200	0.14322	0.019021
6	0.0200	0.14757	0.020115
7	0.0200	0.15142	0.021429
8	0.0300	0.1458	0.02262
9	0.0300	0.14827	0.023859
10	0.0300	0.15054	0.025165
11	0.0300	0.15265	0.02653
12	0.0500	0.14111	0.027703
13	0.0500	0.14241	0.028749
14	0.0500	0.14366	0.029828
15	0.0500	0.14487	0.030941
16	0.0500	0.14604	0.032087
17	0.0500	0.14718	0.033267
18	0.0500	0.14829	0.03448
19	0.0800	0.13397	0.03552
20	0.0800	0.13461	0.036394

Table 13: First moment of g_1 with different values of Q^2: the error on $\Gamma_1^{\rm NS}(Q^2)$ has been added a 100% uncertainty on the low-x extrapolation.

We then proceed following the procedure described in detail in Sect. 4.3 of Ref. [38]: the extraction of couplings is done by combining moments at different values of Q^2 in the chosen range and fitting Eq. (4.6) using MINUIT [39] where g_A and α_s are the chosen as free parameters. The moments at different Q^2 are correlated, since they are computed using the same fitted parametrization. As detailed in Ref. [38] these correlations induce numerical instabilities in the inversion of the correlation matrix and off-diagonal

Proton	$Q^2 = 1$ GeV^2	$Q^2 = 2$ GeV^2
noTMC	0.1410 ± 0.0109	0.1264 ± 0.0083
wTMC	0.1459 ± 0.0116	0.1276 ± 0.0086
Deuteron	$Q^2 = 1$ GeV^2	$Q^2 = 2$ GeV^2
noTMC	0.0493 ± 0.0141	0.0350 ± 0.0058
wTMC	0.0515 ± 0.0158	0.0355 ± 0.0059

Table 14: First moment with and without TMC with $b = 0.2$ for the low-x extrapolation matched at $x = 0.001$.

instabilities due to non-diagonal elements in the correlation matrix becoming dominant. Both these instabilities lead to unreliable results for the extracted couplings.

In order to fix the maximum value of N_{Q^2} for which the extraction of the parameters is numerically stable and reliable, we study the error on the determination of g_A and α_s as we vary the number of included moments. Once we exclude moments with large correlations, we are left with a small number of combinations, which are showed in Tab. 15.

Q^2	g_A	$\alpha_s(M_Z^2)$
2+5	1.04 ± 0.12	0.126 ± 0.005
2+6	1.07 ± 0.16	0.127 ± 0.007
2+7	1.08 ± 0.20	0.128 ± 0.008
2+8	1.02 ± 0.21	0.125 ± 0.012
3+6	1.09 ± 0.13	0.131 ± 0.005
3+7	1.12 ± 0.17	0.132 ± 0.007
3+8	1.04 ± 0.20	0.127 ± 0.013
4+8	1.03 ± 0.17	0.127 ± 0.015

Table 15: Fits with different choices of Q^2.

The combination giving the smallest error ($Q^2 = 2,5$ GeV^2), once we evaluate asymmetric errors, yields

$$\alpha_s(M_Z^2) = 0.126^{+0.004}_{-0.009}\,, \tag{5.11}$$

for the strong coupling, while the error on g_A is found to be symmetric.

The only sources of theoretical uncertainty left to consider are the one due to the choice of factorization scale $Q = k_F m_q$, which we study by varying k_F in the range $0.5 < k_F < 2$ and to the higher–twist contribution. The results of the variation of the factorization scale are shown in Tab. 16. To take into account the higher–twist contribution we take as an estimate the variation of the central values once the lower Q^2 value is moved up to $Q^2 = 3\ GeV^2$ (see Tab.15) and down to $Q^2 = 1\ GeV^2$ ($g_A = 0.99 \pm 0.14$ and $\alpha_s(M_Z^2) = 0.117 \pm 0.007$).

g_A	$\alpha_s(M_Z^2)$	k_F
1.02 ± 0.12	0.126 ± 0.006	0.5
1.04 ± 0.12	0.126 ± 0.005	0.75
1.01 ± 0.11	0.121 ± 0.006	1.5
1.03 ± 0.12	0.120 ± 0.005	2.0

Table 16: Reference fit ($Q^2 = 2,5$ GeV^2, NNLO, 1000 reps, $k_F = 1$) compared with variations of the factorization scale.

In conclusion, we obtain the following result for the determination of the axial coupling g_A and the strong coupling constant α_s

$$\begin{aligned} g_A &= 1.04 \pm 0.12(\text{exp.})^{+0.05}_{-0.06}(\text{theo.}) = 1.04 \pm 0.13(\text{tot.}) \\ \alpha_s(M_Z^2) &= 0.126^{+0.004}_{-0.009}(\text{exp.})^{+0.005}_{-0.011}(\text{theo.}) = 0.126^{+0.006}_{-0.014}(\text{tot.})\,, \end{aligned} \tag{5.12}$$

which are compatible with previous extractions [42,43] from polarized DIS and the Bjorken sum rule.

6. Conclusions

We extracted a parametrization of the spin asymmetries $A_1^{p,d}$, based on all available DIS data using the Monte-Carlo sampling techniques and neural networks as basic interpolation tools. We checked in the process that the statistical methods developed for the unpolarized studies by the NNPDF Collaboration can be naturally extended to handle the new data

sets considered in this work. Our main result is an effective tool, which we used to test different assumptions needed to reconstruct the polarized structure function g_1. As an example of possible applications we compared to previous estimations of experimental sum rules, and we found that the used parametrization for the unpolarized structure function F_2 can be a sizable source of error at low values of Q^2. We also performed a study of the Bjorken sum rule, and the extraction of the axial coupling and the strong coupling, obtaining values which are compatible with previous analysis.

It would be interesting to compare the results obtained for the Bjorken sum rule when determined from global QCD fits to polarized DIS, SDIS and hadron-hadron collisions data, like the one presented in [44], especially once W production data from RHIC will be included in such fits, providing an extra constraints on the light flavours separation.

The present study is also meant to be a first step towards the application of the NNPDF techniques to the determination of a set of polarized PDFs with a faithful error estimation.

Acknowledgments

We thank S. Forte, G. Ridolfi and M. Anselmino for useful discussion and suggestions. We are greatly indebted to R. De Vita (CLAS), L. De Nardo and A. Fantoni (HERMES), A. Bressan (COMPASS) and T. S. Toole (E155) for detailed information on experimental data. The work of LDD is supported by an STFC Advanced Fellowship.

A. Experimental errors

Experimentally, we have

$$A_{||} \simeq \frac{C}{f P_b P_t} \frac{N_- - N_+}{N_- + N_+}\,, \tag{A.1}$$

where

- C is a nuclear correction that depends on the material the target is made of;
- f is the dilution factor which accounts for the fact that only a fraction of the target nucleons is polarizable;
- P_b and P_t are the beam and target polarizations;
- $N_{-(+)}$ is the number of scattered electrons/muons per incident charge for negative (positive) beam helicity.

Thus, most of the errors quoted by experiments are normalization errors.

- EMC [10]: $A_2 = 0$ is assumed; 9.6% overall normalization due to beam and target polarization; multiplicative errors on R and f; additive errors on A_2, the false asymmetry K and the radiative correction.

- SMC98 [11]: $A_2 = 0$ is assumed; multiplicative errors on P_t, P_b, R, f and the polarized background ΔP_{bg}; additive errors on A_2, the false asymmetry ΔA_{false}, the radiative correction and the momentum resolution.

- SMC low-x [12]: $A_2 = 0$ is assumed; multiplicative errors on P_t, P_b, R, f and the polarized background ΔP_{bg}; additive errors on A_2, the false asymmetry ΔA_{false} and the radiative correction.

- E143 [13]: g_2 is evaluated using the Wandzura-Wilczeck relation

$$g_2^{WW}(x,Q^2) = -g_1(x,Q^2) + \int_x^1 \frac{dy}{y} g_1(x,Q^2)\,, \tag{A.2}$$

  using and empirical fit of $g_1/F_1 = ax^\alpha(1+bx+cx^2)(1+Cf(Q^2))$; multiplicative errors on P_t, P_b, f and the nuclear correction C which account for a total 3.7% for the proton and 4.9% for the deuteron; additive uncorrelated error on the radiative corrections.

- E155 [14]: g_2 is evaluated in the same way of E143, but the parameters of the fitted functional form have different values; we will add as a shift the difference between A_1 and g_1/F_1; multiplicative errors on P_t, P_b, f and the nuclear correction C which account for a total 7.6% for the proton; additive uncorrelated error on the radiative corrections.

- HERMES06 [15]: a parametrization for g_2 is fitted to existing data; normalizations errors of 5.2% for the proton and 5% included in the systematic quoted for each data point; additional additive error on the parametrization used for g_2.

- COMPASS [16]: $A_2 = 0$ is assumed; multiplicative errors on P_t, P_b, the dilution factor f and the depolarization factor D; additive errors on the false asymmetry and the radiative correction.

References

[1] S. Forte, L. Garrido, J. I. Latorre and A. Piccione, JHEP **0205** (2002) 062 [arXiv:hep-ph/0204232].

[2] L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione and J. Rojo [NNPDF Collaboration], JHEP **0503**, 080 (2005) [arXiv:hep-ph/0501067].

[3] L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione and J. Rojo [NNPDF Collaboration], JHEP **0703** (2007) 039 [arXiv:hep-ph/0701127].

[4] R. D. Ball *et al.* [NNPDF Collaboration], Nucl. Phys. B **809** (2009) 1 [arXiv:0808.1231 [hep-ph]].

[5] R. D. Ball *et al.* [The NNPDF Collaboration], arXiv:0906.1958 [hep-ph].

[6] J. Blumlein and H. Bottcher, Nucl. Phys. B **636**, 225 (2002) [arXiv:hep-ph/0203155].

[7] J. Blumlein and A. Guffanti, AIP Conf. Proc. **792**, 261 (2005).

[8] M. Anselmino, A. Efremov and E. Leader, Phys. Rept. **261**, 1 (1995) [Erratum-ibid. **281**, 399 (1997)] [arXiv:hep-ph/9501369].

[9] S. E. Kuhn, J. P. Chen and E. Leader, arXiv:0812.3535 [hep-ph].

[10] J. Ashman *et al.* [European Muon Collaboration], Nucl. Phys. B **328** (1989) 1.

[11] B. Adeva *et al.* [Spin Muon Collaboration], Phys. Rev. D **58** (1998) 112001.

[12] B. Adeva *et al.* [Spin Muon Collaboration], Phys. Rev. D **60**, 072004 (1999) [Erratum-ibid. D **62**, 079902 (2000)].

[13] K. Abe *et al.* [E143 collaboration], Phys. Rev. D **58**, 112003 (1998) [arXiv:hep-ph/9802357].

[14] P. L. Anthony *et al.* [E155 Collaboration], Phys. Lett. B **493** (2000) 19 [arXiv:hep-ph/0007248].

[15] A. Airapetian *et al.* [HERMES Collaboration], Phys. Rev. D **75** (2007) 012007 [arXiv:hep-ex/0609039].

[16] E. S. Ageev *et al.* [COMPASS Collaboration], Phys. Lett. B **612** (2005) 154 [arXiv:hep-ex/0501073].

[17] M. Dittmar *et al.*, arXiv:0901.2504 [hep-ph].

[18] G. D'Agostini, *Bayesian reasoning in data analysis: A critical introduction*, World Scientific, 2003.

[19] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford 1995.

[20] K. Hornik, M. Stinchcombe and H. White, *Neural Networks*, vol. 2, 359 (1989).

[21] M. Mitchell, *An introduction to genetic algorithms*, MIT Press, 1998

[22] J. I. Latorre and J. Rojo, JHEP **0401**, 055 (2004).

[23] S. Matsuda and T. Uematsu, Nucl. Phys. B **168** (1980) 181.

[24] A. Piccione and G. Ridolfi, Nucl. Phys. B **513** (1998) 301 [arXiv:hep-ph/9707478].

[25] J. Blumlein and A. Tkabladze, Nucl. Phys. B **553** (1999) 427 [arXiv:hep-ph/9812478].

[26] A. Accardi and W. Melnitchouk, Phys. Lett. B **670** (2008) 114 [arXiv:0808.2397 [hep-ph]].

[27] C. Simolo, arXiv:0807.1501 [hep-ph].

[28] S. A. Larin and J. A. M. Vermaseren, Phys. Lett. B **259** (1991) 345.

[29] S. Wandzura and F. Wilczek, Phys. Lett. B **72** (1977) 195.

[30] L. W. Whitlow, S. Rock, A. Bodek, E. M. Riordan and S. Dasu, Phys. Lett. B **250** (1990) 193.

[31] K. Abe *et al.* [E143 Collaboration], Phys. Lett. B **452**, 194 (1999) [arXiv:hep-ex/9808028].

[32] A. Milsztajn, A. Staude, K. M. Teichert, M. Virchaux and R. Voss, Z. Phys. C **49** (1991) 527.

[33] M. Arneodo *et al.* [New Muon Collaboration.], Phys. Lett. B **364** (1995) 107 [arXiv:hep-ph/9509406].

[34] H. Abramowicz and A. Levy, arXiv:hep-ph/9712415.

[35] W. Buck and F. Gross, Phys. Rev. **D 20**, 2361 (1979), M. Z. Zuilhof and J. A. Tjon, Phys. Rev. **C 22**, 2369 (1980), M. Lacombe et al., ibid. 21, **861** (1980), R. Machleidt et al., Phys. Rep. **149**, 1 (1987)

[36] F. E. Close and R. G. Roberts, Phys. Lett. B **336** (1994) 257 [arXiv:hep-ph/9407204].

[37] R. Abbate and S. Forte, Phys. Rev. D **72** (2005) 117503 [arXiv:hep-ph/0511231].

[38] S. Forte, J. I. Latorre, L. Magnea and A. Piccione, Nucl. Phys. B **643** (2002) 477 [arXiv:hep-ph/0205286].

[39] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).

[40] A. Deur *et al.*, Phys. Rev. Lett. **93** (2004) 212001 [arXiv:hep-ex/0407007].

[41] A. Deur *et al.*, Phys. Rev. D **78** (2008) 032001 [arXiv:0802.3198 [nucl-ex]].

[42] G. Altarelli, R. D. Ball, S. Forte and G. Ridolfi, Nucl. Phys. B **496** (1997) 337 [arXiv:hep-ph/9701289].

[43] G. Altarelli, R. D. Ball, S. Forte and G. Ridolfi, Acta Phys. Polon. B **29** (1998) 1145 [arXiv:hep-ph/9803237].

[44] D. de Florian, R. Sassot, M. Stratmann and W. Vogelsang, arXiv:0904.3821 [hep-ph].